

Full Length Article

Assessing the use of harmonized multisource backscatter data for thematic benthic habitat mapping

Benjamin Misiuk^{a,*}, Myriam Lacharité^b, Craig J. Brown^a^a Department of Oceanography, Dalhousie University, 6299 South St., Halifax, NS, B3H 4R2, Canada^b Institute for Marine and Antarctic Studies, University of Tasmania, Private Bag 129, Hobart, TAS, 7001, Australia

ARTICLE INFO

Keywords:

Benthic habitat mapping
Multibeam echosounder
Backscatter
Seabed mapping
Benthoscape
Object-based image analysis

ABSTRACT

Legacy seabed mapping datasets are increasingly common as the need for detailed seabed information is recognized. Acoustic backscatter data from multibeam echosounders can be a useful surrogate for seabed properties and are commonly used for benthic habitat mapping. Legacy backscatter data, however, are often uncalibrated, rendering measurements relative to a given survey and complicating the use of multisource acoustic datasets for habitat mapping. Recently, 'bulk shift' methods have been proposed to harmonize multisource backscatter layers that overlap spatially, but their application to benthic habitat mapping has not been evaluated. Here, four relative backscatter datasets at the St. Anns Bank Marine Protected Area were harmonized to produce a single continuous surface spanning the extent of available bathymetric data. The harmonized surface was used as a predictor in a benthic habitat ('benthoscape') classification, which was compared to previous results using individual backscatter coverages. Results were similar to those obtained previously, but the harmonized surface provided increased class discrimination, fewer unclassified areas, and predictions that cross dataset boundaries – eliminating the need for manual reclassification by the user. While this generally increases the efficiency and repeatability of the analysis and the useability of the data, we caution that an inappropriate harmonization model is a potential source of error for the classification.

1. Introduction

Acoustic remote sensing of the seafloor has been increasingly adopted over the past three decades for a wide variety of applications, including geological mapping (e.g., Ferrini and Flood, 2006; Hughes Clarke et al., 1996; Misiuk et al., 2018; Plets et al., 2012; Stephens and Diesing, 2014; Todd et al., 1999), marine archaeology (e.g., Passaro et al., 2013; Plets et al., 2011), seafloor environmental change monitoring (e.g., Montereale-Gavazzi et al., 2018; Montereale-Gavazzi et al., 2019; Snellen et al., 2019; van Rein et al., 2011) and benthic habitat studies (e.g., Boswarva et al., 2018; Brown et al., 2012; Kostylev et al., 2003; Lacharité et al., 2018). Multibeam echosounders (MBES) have become the survey tool of choice as costs and availability have recently become less prohibitive (Brown et al., 2011; Brown and Blondel, 2009; Lucieer et al., 2018), and innovation continues with emerging acoustic methods such as interferometric and synthetic aperture sonar (SAS) (Thorsnes et al., 2019). One of the primary benefits of these sonar technologies is their ability to measure both seafloor bathymetry, from which geomorphology can be derived through the production of digital bathymetric models of the

seafloor, and seafloor composition, which can be inferred from the signal of the returning echo – a property referred to as acoustic backscatter intensity (De Falco et al., 2010; Lamarche and Lurton, 2018; Lamarche et al., 2011; Lurton, 2010).

Acoustic backscatter is influenced by seafloor geophysical properties, including sediment grain size, hardness, roughness, and biotic elements (Brown and Blondel, 2009; Lamarche and Lurton, 2018). Sediment-acoustic relationships from MBES are complex due to the influence of angular dependency of the signal intensity caused by ensonification geometry across the MBES swath (Fonseca et al., 2009; Lamarche et al., 2011; Malik, 2019). This is further complicated by the operating frequency of the MBES, and resulting frequency-dependent interaction between the acoustic signal and substrate (Brown et al., 2019). Despite the complexities of signal analysis, progress has been made in developing post-processing methods wherein backscatter can be rendered useful for delineating differences in substrate and habitat (see Lurton and Lamarche, 2015 for a detailed review).

An ongoing challenge of using MBES backscatter for studying the seafloor environment is the lack of standardized and achievable

* Corresponding author.

E-mail addresses: ben.misiuk@dal.ca (B. Misiuk), myriam.lacharite@utas.edu.au (M. Lacharité), craig.brown@dal.ca (C.J. Brown).<https://doi.org/10.1016/j.srs.2021.100015>

Received 23 October 2020; Received in revised form 24 December 2020; Accepted 8 January 2021

Available online 22 January 2021

2666-0172/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

backscatter calibration methods. This is widely acknowledged by the scientific community as a major obstacle (Lamarche and Lurton, 2018; Lucieer et al., 2018), and recent studies have examined backscatter variability and potential calibration approaches – though no single calibration method has achieved widescale adoption thus far (Heaton et al., 2017; Malik et al., 2018; Montereale-Gavazzi et al., 2019; Roche et al., 2018; Weber et al., 2018). Uncalibrated MBES can achieve internally consistent relative backscatter measurements when acquisition parameters are constant within a survey, but multiple surveys using uncalibrated MBES systems are highly likely to produce inconsistent backscatter datasets. When a dataset comprises backscatter from multiple sources (e.g., different MBES systems, time periods, operating frequencies, etc), it becomes difficult to establish robust relationships between the backscatter signal and substrate properties, hindering the use of backscatter for seabed mapping applications. As seabed mapping approaches become increasingly automated, it is critical that predictor variables consistently represent environmental characteristics across the extent of the mapped area, and the use of uncalibrated multisource backscatter datasets can invalidate this requirement.

Multisource uncalibrated backscatter datasets may occur for several reasons. They are often produced when mapping extensive or complex areas that require multiple surveys to complete. While surveys by multiple vessels or MBES systems are almost certain to produce incompatible datasets, incompatibility can also arise from surveys by the same vessel and MBES system collected at different time periods (Lamarche and Lurton, 2018). Opportunistic data collected for other purposes may also be used to fill data gaps, yet these are unlikely to avail of any calibration procedures. Further, the incorporation of legacy data for seafloor mapping can be particularly challenging when acquisition and processing parameters are unknown or outdated, and the raw data are not available for reprocessing. These are difficult issues, yet developing solutions to utilize multisource backscatter datasets is critical for maximizing the potential of contemporary and legacy seafloor mapping datasets.

One potential approach to handling relative backscatter inconsistencies is to apply harmonization methods to the post-processed gridded backscatter mosaics. Misiuk et al. (2020) demonstrated backscatter harmonization that relies on areas of mutual survey overlap to calibrate disparate datasets – termed “bulk shift” approaches. These may be compared to the use of natural reference areas for system calibration (Roche et al., 2018; Weber et al., 2018), but with *ad hoc* areas of mutual survey coverage from raster grids rather than pre-defined calibration patches. While undoubtedly less rigorous, bulk shift approaches do not require the raw data or a pre-planned calibration procedure, making them potentially useable for harmonizing datasets from a variety of origins, including opportunistic and legacy sources. Furthermore, the use of gridded data implies aggregation, which may compensate for differences in fine-scale substrate characterisation between sensors.

The implications of bulk shift harmonization approaches for benthic habitat mapping – for which a single consistent backscatter layer is highly desirable – are promising. For many statistical approaches to benthoscape classification and species distribution modelling (SDM), there exist a host of statistical and practical difficulties associated with deriving spatially continuous habitat predictions from multiple uncalibrated backscatter datasets. These include differences in ground truth sampling effort between datasets, discontinuous characterization of other environmental gradients across datasets (e.g., morphology, oceanography), and abrupt disagreement of segmented model predictions at dataset boundaries (Lacharité et al., 2018). Use of the bulk shift backscatter harmonization for habitat mapping is therefore motivated by a need for continuous and internally consistent surrogates for seabed substrate properties, which should strive to match the extent and quality of other common morphological and oceanographic variables. For geospatial modelling, backscatter layers should ideally be robust to differences in sampling effort and dataset boundary effects.

The effectiveness of bulk shift harmonization methods for producing continuous and consistent substrate surrogates for benthic habitat

mapping has not yet been demonstrated. It is necessary to investigate whether bulk shift harmonization can alleviate issues associated with ground truth fragmentation and dataset discontinuity, but several well-recognized challenges for this approach must also be considered. Differences between data acquisition and processing parameters (sonar system, operating frequency, beam geometry) present potential challenges for backscatter harmonization, and the impacts of these challenges on habitat mapping methods and products is currently not resolved.

In this study, we investigate the applicability of bulk shift backscatter harmonization to benthic habitat classification through a comparative study using individual and harmonized backscatter layers. Lacharité et al. (2018) derived a benthic habitat map for a portion of the St. Anns Bank Marine Protected Area (MPA), off the coast of Nova Scotia, Canada, by using four separate MBES datasets collected over multiple years, with different MBES systems, using different operating frequencies. This study employed object-based image analysis (OBIA) – a popular method to define thematic seafloor units (e.g., Diesing et al., 2014; Innangi et al., 2019; Janowski et al., 2020; Lucieer, 2008) – to combine results from the four coverages into a single seamless habitat map. Clear discontinuities at dataset boundaries and uneven spatial distribution of ground truth data across the MBES datasets were recognized difficulties in this study, requiring manual reclassification to correct. Because each MBES coverage overlaps with adjacent coverages, it is possible to use bulk shift approaches to generate a harmonized backscatter layer for the habitat classification – results from which can be compared to those generated from the individual coverages.

The goals of this study are:

- i) to generate a harmonized backscatter mosaic using legacy datasets for St. Anns Bank (Nova Scotia, Canada) and investigate internal methods of validation;
- ii) to use the harmonized backscatter mosaic to produce a benthic habitat classification (benthoscape) map and compare results with those generated using individual coverages by Lacharité et al. (2018).

2. Methods

2.1. Study area and data sets

St. Anns Bank is located off the eastern coast of Nova Scotia (Atlantic Canada), approximately 16 km east of Cape Breton Island (Fig. 1). The MBES data cover an area of 2870 km² ranging in depth from ~20 to 275 m and extending from the bank to deeper waters in the adjacent Laurentian Channel. The study area is located within the St. Anns Bank MPA, designated in 2018 under Canada's Oceans Act to conserve local ecological diversity.

Four surveys were conducted in the area using different MBES systems. The Canadian Hydrographic Service (CHS) first surveyed portions of St. Anns Bank in 2010 with the CCGS *Matthew*, CSL *Plover*, and CSL *Pipit*. The *Matthew* mapped the largest portion of the area using a Kongsberg Simrad EM710 70–100 kHz MBES; the *Plover* and *Pipit* mapped a smaller shallow area in the western part of the site, each using Kongsberg Simrad EM3002 300 kHz MBES systems. Further surveys were conducted aboard the CCGS *Creed* in 2012 at the western and southern parts of the study area using a Kongsberg Simrad EM1002 95 kHz MBES. Finally, acquisition of a fourth dataset at the northernmost part of the study area was contracted by Fisheries and Oceans Canada to a private survey company in 2013 and was completed by the M/V *Dominion Victory* using a pole mounted Reson7111 100 kHz MBES.

The raw backscatter signal from each MBES was recorded during acquisition and subsequently processed using the QPS software suite. The Fledermaus Geocoder Toolbox (FMGT) was used to extract the backscatter as time series snippets from the raw MBES data. Radiometric corrections were applied using default settings within the software, followed by angle-varying gain and anti-aliasing. The resulting dataset was aggregated and gridded at a 5 m horizontal resolution (WGS84 UTM Zone 21 N).

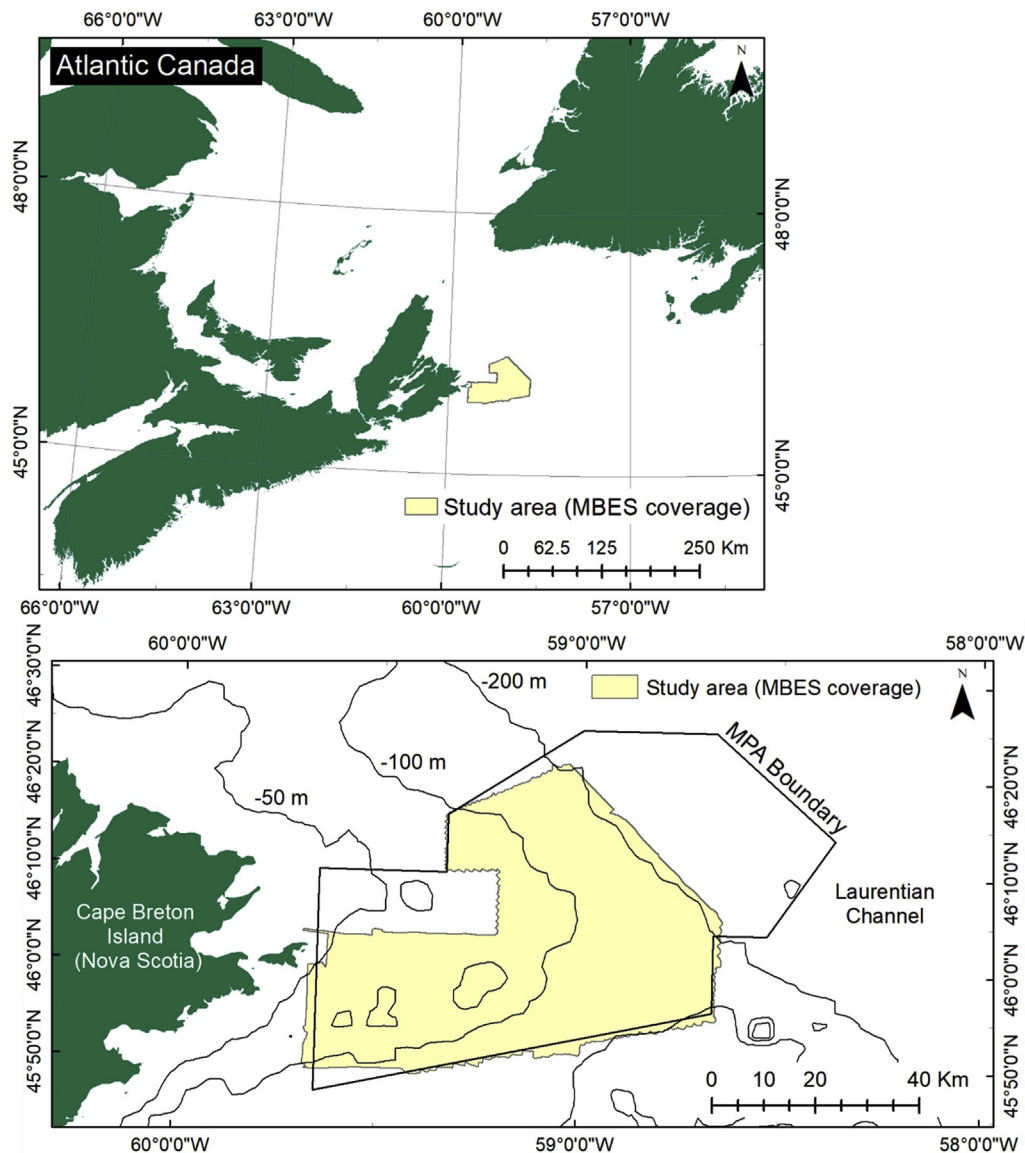


Fig. 1. Location of the study site within the St. Anns Bank MPA off the coast of Nova Scotia, Canada.

A total of 4214 images at 61 stations distributed throughout the study area were included in this analysis (Fig. 2). Seabed photographic surveys were conducted in September 2014 (14 stations) from the R/V *Strait Explorer* and in November 2013 (33 stations) from the R/V *Strait Signet*. A passive dropdown camera system was deployed at each station at an altitude of 1–2 m above the seafloor along a linear transect (200–500 m in length), yielding between 34 and 110 usable photographs per station. The drop-camera system was fitted with a 10 Megapixel Kongsberg OE14-408 Underwater Digital Stills Camera with integrated lasers (16 cm apart) and video cameras. Photographs were taken based on the real-time video feed aboard the ship, and seabed positioning was estimated by calculating offsets from geolocation acquired at the surface.

Imagery from three previous federal government surveys was also incorporated into the dataset. Three stations were included from a 2009 survey aboard the CCGS *Hudson* using a drop-camera (4KCam - Geological Survey of Canada), which housed a Canon Rebel Eos Ti 12-megapixel camera triggered automatically when the system touches the seafloor. Two stations were also included that were acquired using the drop camera system Campod (Fisheries and Oceans Canada), which consists of a downward-facing Nikon D300 12-megapixel camera mounted on a tripod. Additional imagery from the 2010 CCGS *Matthew* cruise

comprised nine stations acquired using a Natural Resources Canada system (Deep Imager) supporting a Sony 520CX HD camera with lights.

2.2. Resampling and harmonization

To match the 50 m horizontal grid resolution used in Lacharité et al. (2018), all backscatter datasets were resampled from the initial 5 m raster grid using bilinear interpolation. The EM710 grid was the most extensive and overlapped all other datasets; each other grid was aligned with it during resampling. The resampled backscatter datasets were then harmonized using the bulk shift method described by Misiuk et al. (2020), with the EM710 dataset as the *target* layer. Each dataset was first corrected with reference to the EM710 data, then all corrected datasets were mosaicked to produce the harmonized backscatter layer.

The methodology employed to correct backscatter datasets to the target dataset is detailed in Misiuk et al. (2020) and is summarized here. This approach models corrected backscatter values using differences between overlapping *shift* and *target* datasets, and optional associated bathymetric variables. First, the values of the EM710 data and the dataset being corrected (the *shift* layer) are extracted from the area of overlap. The values of the *shift* dataset are subtracted from the *target* data (EM710)

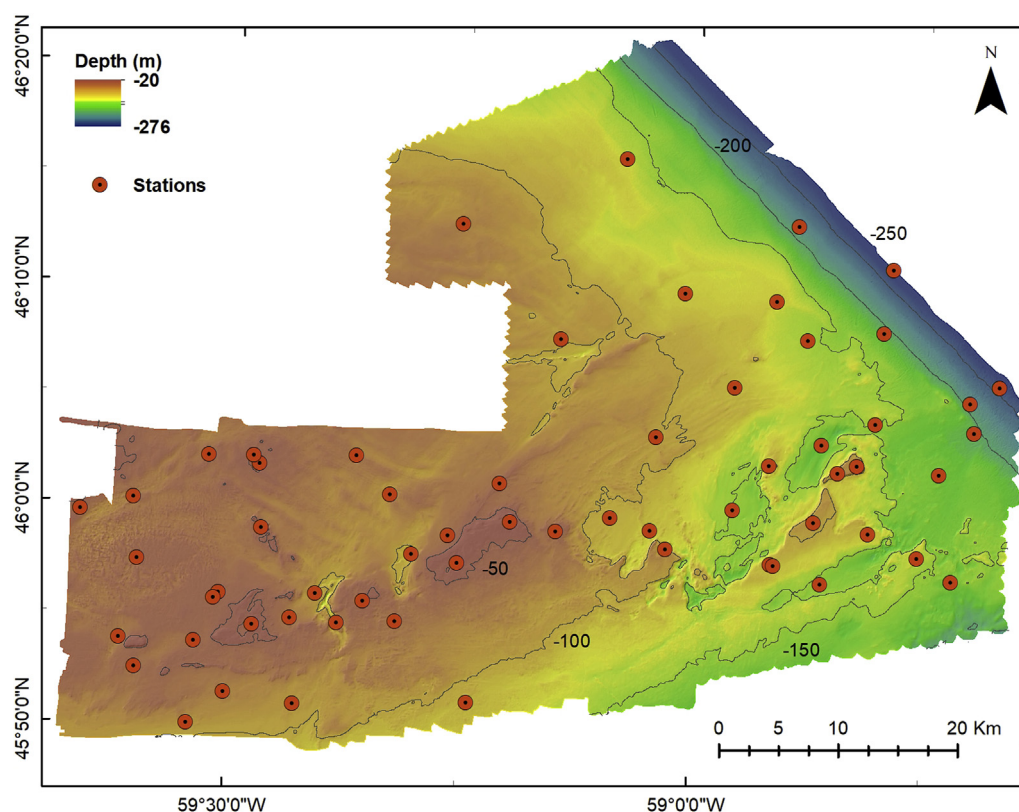


Fig. 2. Bathymetry at St. Anns Bank with 50 m isobaths and underwater camera station locations.

to produce error measurements for each raster cell at the overlap. The error is treated as a response variable, and a model is fit using the dB level of the *shift* layer as a predictor, along with optional additional variables such as water depth, which was selected for its previous utility in reducing harmonization error (Misiuk et al., 2020). The model is then used to predict the error across the remainder of the *shift* dataset, outside the area of overlap. In other words, the difference between the two datasets that is observed where they overlap is used to infer, in general, how much the *shift* dataset must be adjusted by to match the *target* as closely as possible. This model is used to generate values that are added to the original data to obtain the desired corrections. Given a modelling method that is more complex than a simple mean shift, this procedure simultaneously rescales the *shift* data to the *target* and corrects for non-uniform differences in backscatter calibration between datasets. Finally, the model fit is assessed using plots of the relationships between predictor variables and the response (dB error), residual plots, and statistical distribution plots and statistics. While any regression modelling method can be used within this framework to predict the error of the *shift* dataset and correct it with reference to the *target*, previous simulations have suggested the preference of low variance models (Misiuk et al., 2020).

Table 1
Benthoscape classes defined for St. Anns Bank (Atlantic Canada).

Class	Biophysical characteristics	No. images	Percentage of images
A	Mud	813	19%
Asp	Mud with seapens	243	6%
B	Gravelly sand/mud; <50% cobbles/gravel	524	12%
C	Till >50% cobbles/gravel	1115	26%
D	Till with coralline algae	956	23%
E	Gravel with crinoids	216	5%
F	Sand with sand dollars	347	8%

All corrections here were applied using the ‘bulkshift’ function (available in Misiuk et al., 2020) in R 3.6.3 (R Core Team, 2019). Multiple regression was selected for the error models, and ‘depth’ from the bathymetric raster (Fig. 2) was supplied as a predictor variable in addition to the *shift* dB values. Model fit was assessed using error and residual plots, probability density functions (PDFs), and the mean absolute error (MAE) between the *target* dataset and the layer being shifted. The harmonization quality was also assessed visually in the final backscatter mosaic.

2.3. Classification

Segmentation and classification of the bathymetry and harmonized backscatter mosaics was conducted according to methods in Lacharité et al. (2018) using the software eCognition v9.5. Segmentation and supervised classification were originally performed separately on the four multibeam coverages. A portion of the classified image-objects were subsequently manually re-classified to produce a single seamless benthoscape map at St. Anns Bank by correcting for obvious classification errors, filling gaps due to sparse ground-truthing data in select regions, and ensuring continuity across coverage boundaries. Ground-truth data used in Lacharité et al. (2018) comprised 4214 benthic images, each classified to one of seven benthoscape classes (Table 1), located at 61 stations across St. Anns Bank (Fig. 2). The same ground-truth dataset was used here.

Mosaics were segmented using the Multiresolution Segmentation (default settings; scale parameter: 15, shape: 0.1, compactness: 0.5) within eCognition, with acoustic backscatter strength given twice the weight of bathymetry. Classification of resulting image-objects was performed using a nearest neighbor supervised scheme based on the seven benthoscape classes. While only the EM710 coverage was used for defining classes in the original study, the harmonized backscatter mosaic enabled class definition here using all image-objects containing ground-

truth. Samples were selected by overlaying georeferenced classified images over image-objects. In most cases (80%), a single benthoscape class was observed at individual stations. If not, when suitable, adjacent image-objects were also classified to represent the full local variability of benthic classes. Class definitions consisted of five variables within image-objects: mean and standard deviation of bathymetry and backscatter intensity, and mean bathymetric position index (BPI; radius of 500 m) – the same suite of variables used by Lacharité et al. (2018). BPI was derived with the Benthic Terrain Modeller Toolbox v3.0 (Walbridge et al., 2018) in ArcGIS. The classification assigns a membership value [0, 1] to each image-object, representing the Euclidean distance to each class center, after standardization. Image-objects were then classified according to their highest membership values (excluding samples), with those of membership values < 0.1 remaining ‘unclassified’.

2.4. Benthoscape map comparison

The classification resulting from the harmonized backscatter mosaic was compared to the benthoscape maps generated by Lacharité et al. (2018). For comparison, the original map from the previous study (*before* manual re-classification) is termed ‘Original’, the final map from the

previous study (*after* manual re-classification) is termed ‘Manual’, and the benthoscape map produced here is termed ‘Harmonized Backscatter’.

Comparison with the Original and Manual maps from Lacharité et al. (2018) were conducted using two metrics: 1) membership values to the assigned classes and spatial patterns in gain (or decrease) in these values for all image-objects, and 2) changes in class assignment by proportion of surface area. Dataset boundaries were also manually examined for continuity to qualify the effectiveness of the harmonization.

3. Results

3.1. Harmonization

Visual analysis suggested the bulk shift procedure was effective at rescaling the three datasets with respect to the EM710 backscatter values (Fig. 3). Few dataset compilation artefacts remained after harmonization, even at the boundary of the Reson7111 survey, which was originally of disparate scale, with arbitrary units ranging between ~130 and 180. The bulk shift procedure effectively normalizes these data to a scale comparable to the relative dB scale of the *target* dataset (EM710). PDFs illustrate the change in data distribution of the *shift* layers at the area of overlap

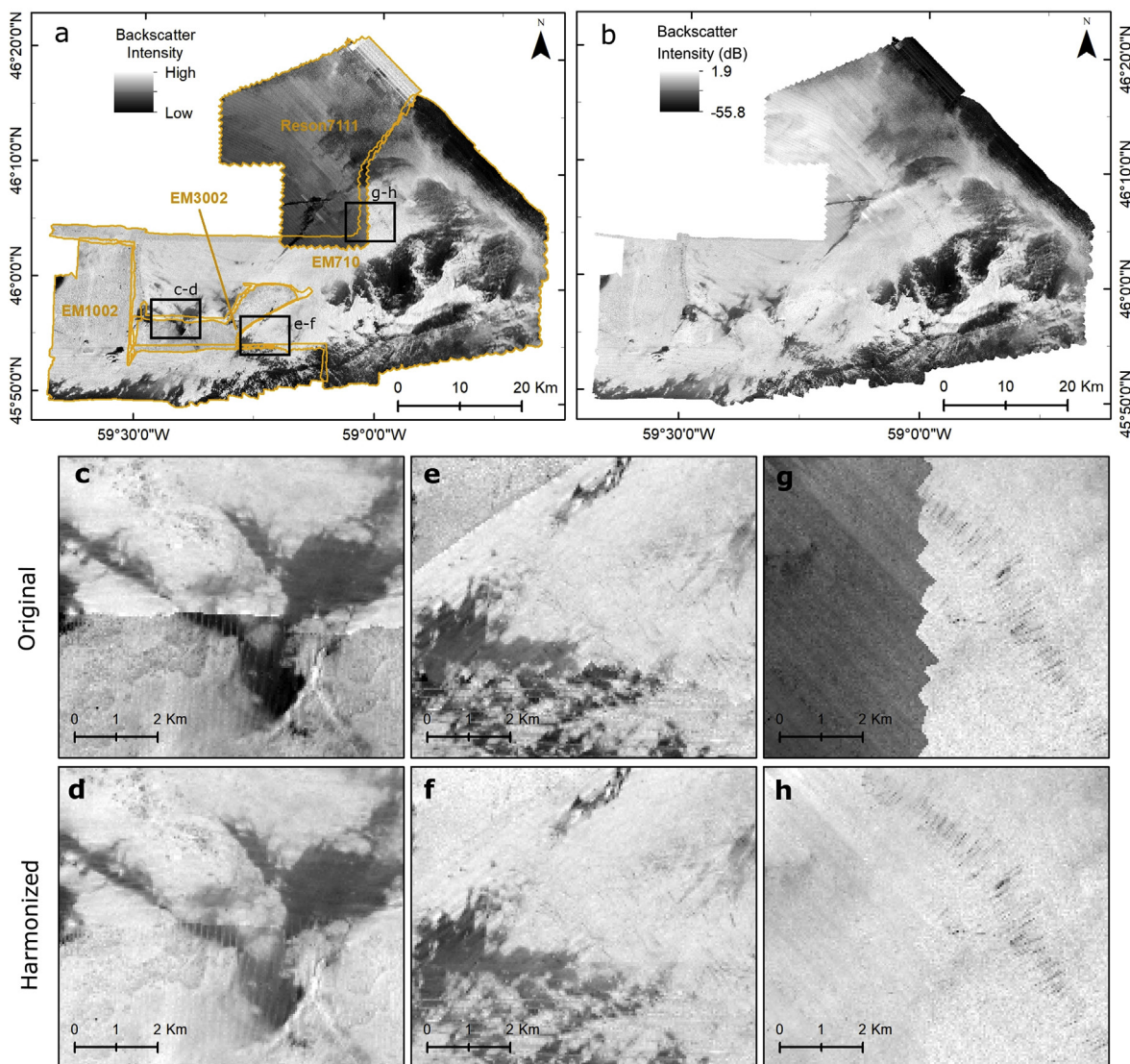


Fig. 3. Results of backscatter dataset harmonization at the St. Anns Bank MPA (Nova Scotia, Canada), with (a) backscatter datasets from four separate surveys, and (b) the harmonized backscatter mosaic after corrections using the bulk shift approach (Misiuk et al., 2020). Dataset boundaries are compared before and after harmonization for: (c–d) EM710 and EM3002 datasets; (e–f) EM710, EM3002, and EM1002 datasets; (g–h) EM710 and Reson7111 datasets.

(Fig. 4). The similarity of data values and distributions to the EM710 data at the overlaps is increased substantially by the bulk shifts, with initial MAE values of 8.12, 7.58, and 167.13 dB reduced to 1.62, 1.18, and 1.38 dB after harmonization for EM3002, EM1002, and Reson7111 datasets, respectively.

Linear relationships were observed between backscatter error and depth at the areas of overlap (Fig. 5). The slope of the relationship was gradual for EM3002 and EM1002 datasets (0.02 and 0.01 dB/m), but error increased more rapidly with depth for the Reson7111 dataset (0.13 dB/m). The overlap between Reson7111 and EM710 surveys also spanned a greater depth range than the other surveys – from 75 to 255 m. Although the relationships between depth and error appeared linear, interpretation of the statistical significance of the multiple regression coefficients is avoided, as the spatial autocorrelation of proximal raster pixels may compromise their independence. Residuals do, however, appear to largely conform to a normal distribution for all models (Appendix; Figure A1) – though the presence of outliers is notable (Fig. 5).

3.2. Classification

Segmentation using the harmonized backscatter mosaic yielded 2074 image-objects – slightly fewer than the combined 2189 image-objects in the segmentation of the four individual coverages in Lacharité et al. (2018). Classification of these image-objects into the seven benthoscape classes produced a map similar to both the Original and Manual maps (Fig. 6). Local variability in benthoscape classes was more pronounced here than in the final benthoscape map ('Manual') in the previous study, but interestingly, was similar to the Original map derived from the four individual coverages. Thirty image-objects were 'unclassified' in the Harmonized Backscatter map (1.4% of image-objects; 0.4% of surface area), while 87 were 'unclassified' among individual coverages in the Original map (4% of image-objects). This unclassified area (0.4% of total surface area) of the Harmonized Backscatter map could indicate the presence of un-sampled benthoscape class(es) and/or acoustic mapping errors or outliers.

Excluding unclassified image-objects in both the previous and new maps (2.4% of the surface area), the relative proportion of benthoscape classes was similar between Original, Manual, and Harmonized Backscatter maps, with few exceptions. Classification at 71.4% of the surface area was consistent with previous maps; the total agreement across all three maps was 63.4%, while agreement between only the Manual and Harmonized Backscatter (but not the Original) maps was 6.0%. Agreement between only Original and Harmonized Backscatter (but not Manual) maps was 2.0%. Where the class of the Harmonized Backscatter map differed from previous maps (28.6% of surface area), the Original and Manual maps agreed at 27.4% of the surface area. True confusion in classification occurred at 1.3% of the surface area, where all three maps

disagreed. Changes in relative cover of benthic classes in the new map vary between $\pm 8\%$ of surface area relative to previous maps (Fig. 7). Coverage of benthoscape class B increased most, replacing the benthoscape classes A, C and F.

Membership values of individual image-objects increased when using the Harmonized Mosaic (average: 0.82 excluding image-objects used as samples, but including unclassified image-objects with a value of 0) relative to the average of the initial segmentation and classification among the four coverages in the previous study (0.73; Fig. 8). Positive gain in membership values were observed at the boundaries of the four initial coverages, especially within the Reson7111 coverage, where many unclassified image-objects occurred in the Original map, and at the EM3002 boundary with the EM710 coverage (Fig. 9). Low membership values were attributed to the lack of ground-truth data of benthoscape classes in the original study, which were obvious continuations of adjacent coverages.

4. Discussion

4.1. Harmonization

Visual and statistical analysis of the St. Anns Bank backscatter mosaic suggested that harmonization was effective for relative calibration of these datasets. Discontinuities at the dataset boundaries are the most obvious visual indicators of poor harmonization quality, and continuity between MBES coverages was clearly enhanced in the harmonized backscatter mosaic (Fig. 3). Though it may not be surprising that common substrate types (or backscatter levels) of overlapping datasets align well after a bulk shift, rarer substrate types and sporadic features may be useful for diagnosing the transferability of corrections to the tails of the dataset. Obvious discontinuities across survey boundaries at distinct seabed features may indicate that the harmonization model may not be appropriate for the entire distribution of data values. The mean absolute error across all overlapping grid cells decreased substantially after the bulk shift, suggesting a useful model, yet it is important to recall that fit statistics of the training data can suffer from over-optimism regarding model transferability. This may motivate the selection of "low variance" modelling approaches, such as generalized linear models, for their transferability, rather than more complex, higher variance models (Misiuk et al., 2020).

Although it is tempting to rely on the significance estimates of model coefficients to determine the quality of a model for a bulk shift of backscatter values, these are likely to suffer from non-independence caused by the spatial autocorrelation of grid cells. It is still critical to diagnose the model fit quality though, and this can be accomplished using tools such as descriptive plots of predictive relationships (e.g., univariate or partial dependence plots) and model residuals – yet care

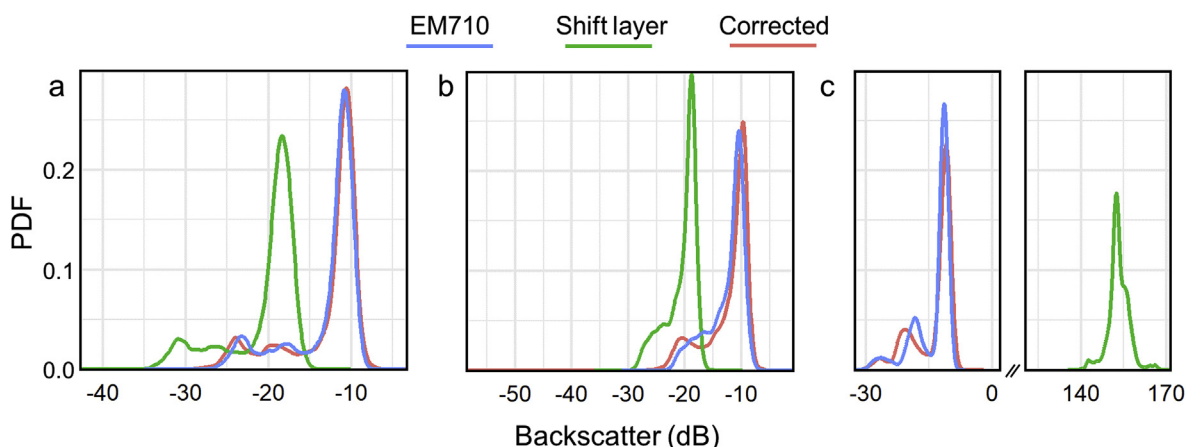


Fig. 4. PDFs of the EM710 target layer, compared to the A) EM3002, B) EM1002, and C) Reson7111 shift layers before and after correction, at the areas of overlap.

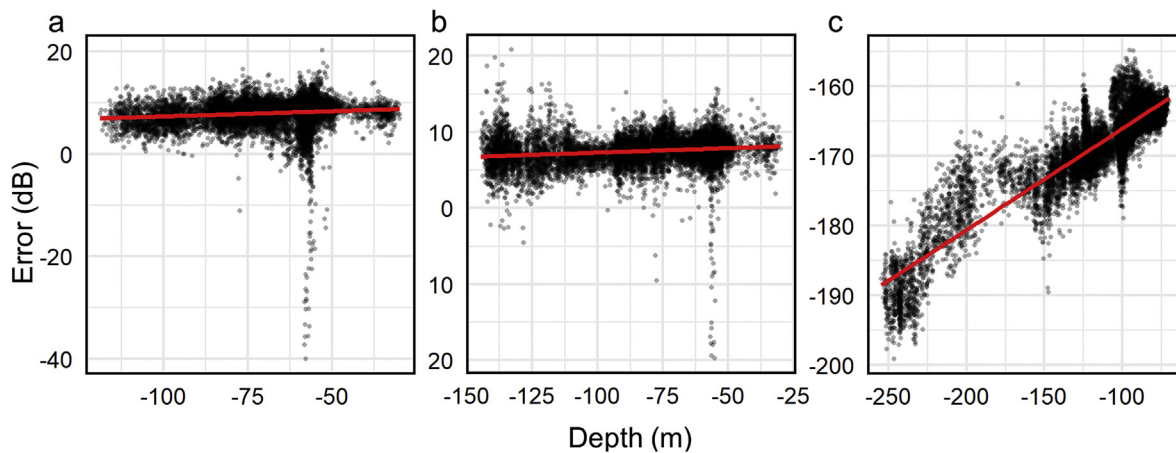


Fig. 5. Error between EM710 backscatter values as a function of depth at areas of overlap for A) EM3002, B) EM1002, and C) Reson7111 datasets. Note that the line of best fit is a univariate regression between depth and backscatter error for visualization and does not represent the exact coefficients used in the full multiple regression models.

must be taken in assigning statistical significance to such relationships. Alternatively, a more sophisticated model validation could be performed to account for autocorrelation. Although results from Misiuk et al. (2020) suggested that visual interpretation of the harmonized mosaics aligned well with test statistics in simulation, and that multiple regression models largely avoided overfitting the data at the area of survey overlap, it is still necessary to obtain a well-rounded understanding of the model quality using a diverse set of tools that include visual assessment of the mosaic, and visual and statistical assessment of the model error and residuals. This is particularly important for deciding whether to include additional predictors such as water depth for reducing the error between backscatter datasets. Simulations in Misiuk et al. (2020) suggested that water depth information was useful for reducing error in many of the models, and a potential interaction between the depth and substrate-induced signal attenuation was hypothesized. Where additional predictors are included, statistical relationships with the backscatter error must be carefully evaluated.

It is also interesting to consider model statistics and assessment in the context of the inherent variability (noise) of backscatter data, which commonly yields a ‘speckled’ appearance when mosaicked as a raster. The model intercept is expected to usually explain most of the error between backscatter datasets, especially where their scales are drastically different, yet the natural variability within backscatter datasets may account for most or all of the residual error after accounting for the intercept. The simplest form of bulk shift is the addition of a single offset value that corrects the mean of the error (i.e., residuals) to zero. Having zeroed the mean of the residuals though, further model coefficients have the difficult task of attempting to separate systematic error from noise, and this yields a high potential for overfitting using flexible models (Misiuk et al., 2020). Given high data variability, accounting for the slope of the error via linear regression may allow for additional modest decreases in error while largely avoiding the influence of backscatter data variability. In this study, the inclusion of even a small slope term for the depth variable (~ 0.02 dB/m) predicts a ~ 2 dB offset in error between the deepest and shallowest areas for the EM1002 and EM3002 datasets (Fig. 4), and this relationship appears fairly consistent amidst highly variable data.

The Reson7111 dataset displayed greater internal variability and depth variability than the Kongsberg datasets. In addition to minor apparent differences in the backscatter levels of adjacent track lines (Fig. 3), there appeared to be a loss of signal quality in the easternmost section of the Reson dataset, where depth increases to > 200 m (Fig. 2). The plot of depth vs. error (Fig. 5c) suggested a strong relationship (i.e., increasing error with depth), which was also observed previously (Misiuk et al., 2020). Abrupt inconsistencies between adjacent survey lines, and

in deeper waters, may be caused by changes to operating parameters of the echosounder during acquisition (e.g., changes to gain). This can occur manually or automatically in Reson systems, which can help to produce higher quality bathymetric data, but at the expense of the backscatter data quality (Lurton and Lamarche, 2015). The addition of bathymetry as a predictor appears to have ameliorated some of the depth-dependent error compared to the EM710 data, which may be a proxy for depth-dependent changes to the sonar operating parameters. Despite the depth inconsistency and internal variability though, the average error between this dataset and the EM710 was reduced to levels comparable to the other datasets after bulk shift harmonization.

The decision to harmonize multisource backscatter data should be purpose-driven, as it requires assumptions regarding the compatibility of the datasets that can introduce additional uncertainty into the estimation of substrate properties. Conceptually, backscatter mosaics are a model of seafloor reflectivity based on a number of oceanographic, substrate, and operating conditions (Lamarche and Lurton, 2018). The harmonization of backscatter mosaics, therefore, relies on modelling a set of models, and the user should be conscious of the possibility of error propagation (i.e., the accumulation of error through multiple stages of modelling). The consequences of error propagation may depend on the intended purpose of the backscatter layer – potentially differing, for example, between applications such as sample planning and high-resolution surficial geology mapping. In the context of thematic mapping, the potential impacts of error propagation should be contextualized given the level of thematic and spatial map units. This is closely related to the analysis scale; low resolution backscatter mosaics may represent either a) broad aggregations of higher-density soundings, or b) a large beam footprint, and may therefore already constitute a coarse approximation of reflectivity over an extensive area. Concerns over the propagation of error may be correspondingly lower at broad scales compared to finer scale applications and mapping units. Regardless of the scale of analysis though, a critical assumption when implementing a bulk shift harmonization is that each backscatter dataset is internally consistent (Misiuk et al., 2020), implying that relationships observed at one part of the dataset (i.e., at the area of overlap) can be transferred to the remainder.

4.2. Classification

The classification generated using the harmonized backscatter mosaic in this study produced similar results to those in Lacharité et al. (2018), providing confidence that the harmonized mosaic is a consistent and useful proxy for deriving a thematic seafloor map. Despite similar outcomes though, use of a harmonized backscatter layer can provide several benefits for classification. Accuracy can be enhanced through improved

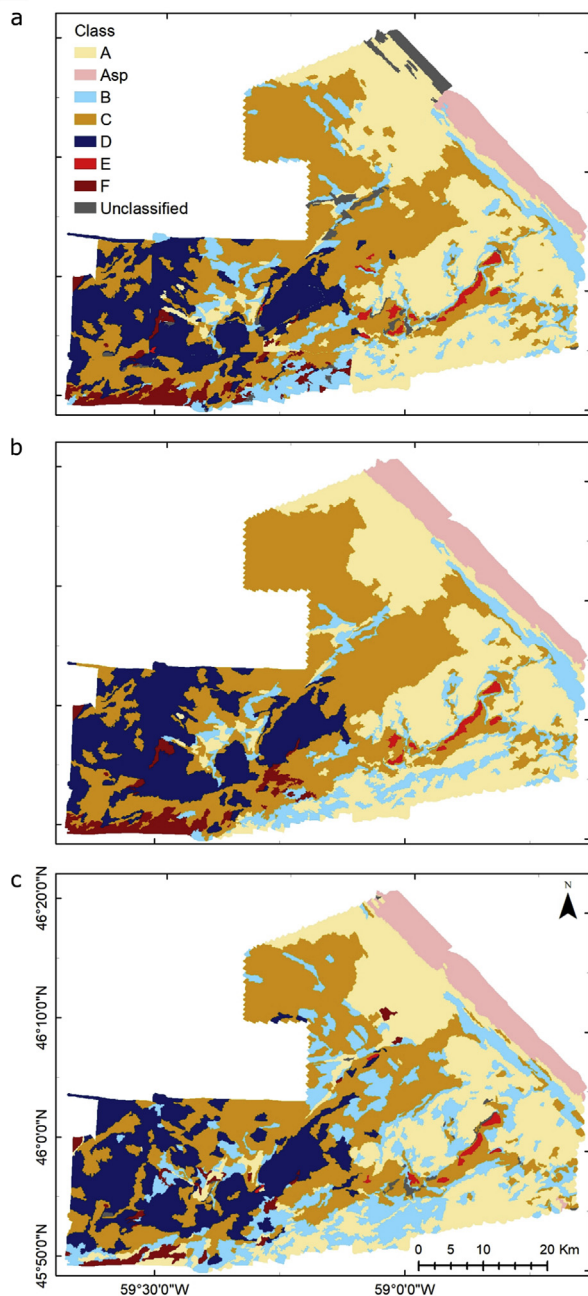


Fig. 6. Benthoscape maps resulting from (a) initial segmentation among all four separate coverages ('Original'; Lacharité et al., 2018), (b) after manual re-classification ('Manual'; Lacharité et al., 2018), and (c) using the new Harmonized Mosaic (this study).

class definitions based on more ground-truth samples. In this study, the average membership value increased relative to the Original thematic map, and fewer image-objects were 'unclassified'. Gains in membership value occurred where ground-truth samples were sparse (e.g., Reson7111 dataset) and at boundaries between datasets, where uncertainty was greatest. We note that the greatest gains in membership were obtained at several clusters of high uncertainty identified in the original analysis. Such gains in confidence would be unexpected if the backscatter layer used for classification was not consistent across the area of ground-truth coverage. Furthermore, predicted classes crossed dataset boundaries where they previously caused classification discontinuities (prior to manual reclassification). The increased availability of training samples enabled by the use of a harmonized layer could also allow for assessment

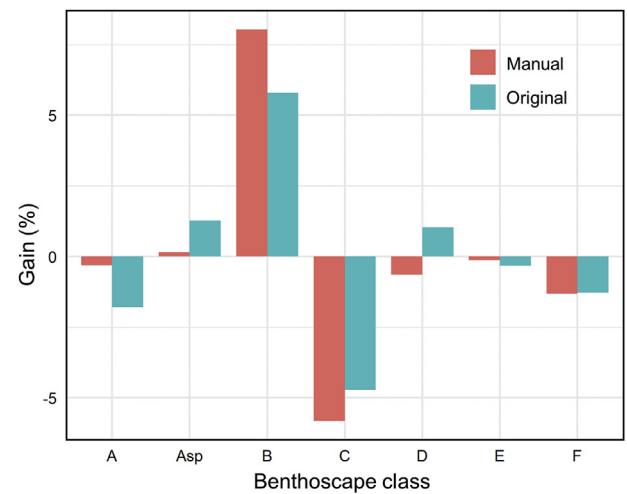


Fig. 7. Gain or loss in surface area of each benthoscape in the new map using the Harmonized Mosaic from this study relative to initial classification of the four separate coverages ('Original'; Lacharité et al., 2018) and after manual re-classification ('Manual'; Lacharité et al., 2018).

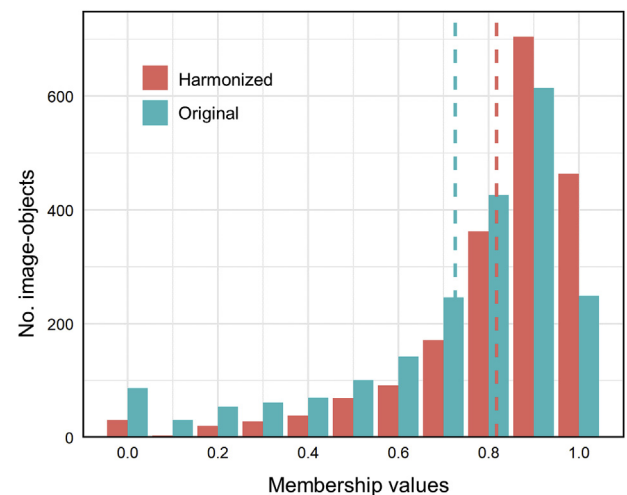


Fig. 8. Membership values of all image-objects in the Harmonized Backscatter map ($n = 1981$) and Original separate coverages ($n = 2080$; Lacharité et al., 2018). Averages of both groups are shown with dashed lines.

of the classification accuracy (e.g., with cross-validation), which was infeasible in the analysis conducted on individual datasets. Practically, using a single harmonized backscatter layer streamlines the analysis, decreasing the need to manually re-classify image-objects.

Although the use of a single harmonized backscatter layer is efficient and desirable for some applications, separate analyses on individual multisource datasets may be preferable when they have been adequately ground-truthed. Different operating specifications of individual sonar systems can ultimately cause them to detect different seabed properties, with differences in penetration depth relating to the operating frequency being among the greatest concerns (Brown et al., 2019; Gaida et al., 2018, 2019; Hamilton, 1972). While subtle differences in mapped seabed characteristics between similar sonar systems may be moderated by data aggregation (e.g., gridding backscatter measurements) and the scale of analysis (e.g., classification of broad-scale image-objects), we expect such resilience to diminish with increasing difference between acquisition parameters (e.g., operating frequency). This was supported by findings in Misiuk et al. (2020), in which increases in backscatter harmonization error were observed with increasing difference between operating

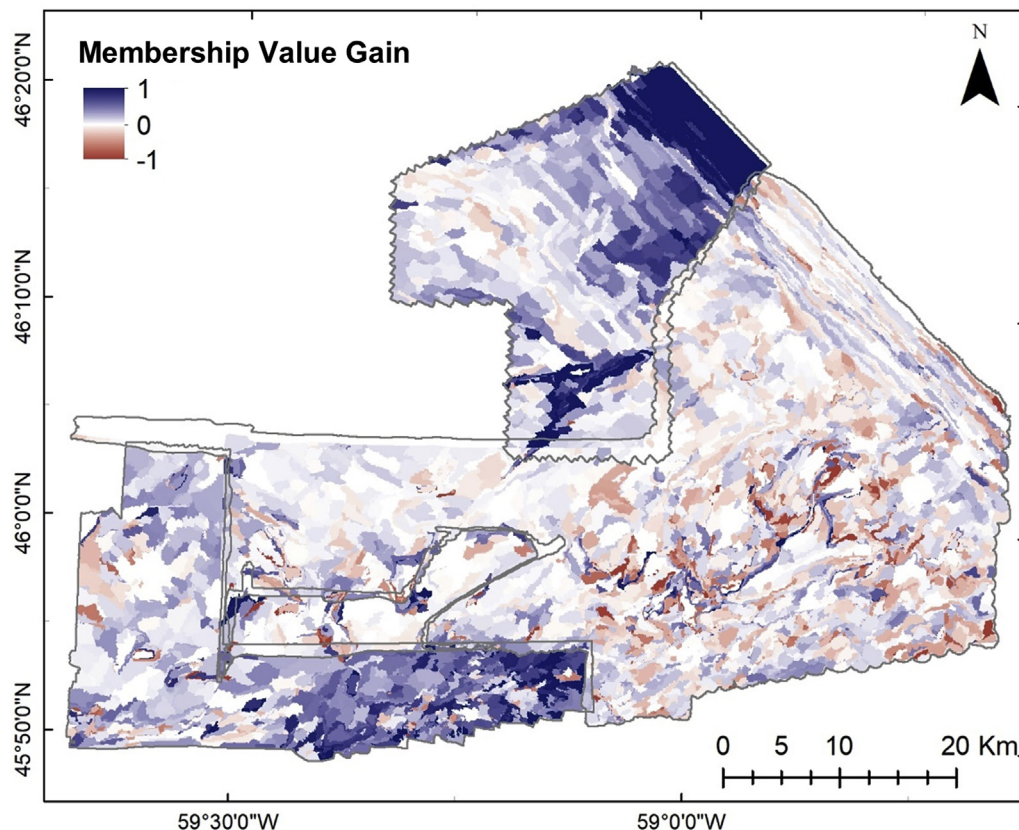


Fig. 9. Gain in membership value [-1,1] of the new map derived from the Harmonized Backscatter mosaic relative to the Original segmentation and classification in Lacharité et al. (2018). Extent of individual coverages is shown in dark grey. Image-objects used as samples are excluded.

frequencies. Given adequate sampling intensity and distribution, individual and separate classifications for each MBES dataset require fewer assumptions regarding the transferability of information between datasets, and separate classifications may therefore be preferable.

4.3. Other applications and context

The focus of this paper is on the applicability of backscatter harmonization methods for benthoscape mapping using multisource sonar data, yet there are other applications for which harmonization may be useful. Benthic species distribution modelling (SDM) requires that the presence or abundance of an organism be sampled across a range of environmental variables to model and predict its distribution. Backscatter data can provide essential substrate information for modelling benthic species habitat, yet where multisource acoustic datasets are compiled to provide environmental data, internally relative backscatter datasets may fragment sampling of broader environmental variable ranges. While this essentially presents similar difficulties as are addressed in this study, the modelling of single taxa presents a continuous (i.e., regression) rather than discrete (i.e., classification) problem, where increased sample size and appropriate distribution are almost always critical (Hernandez et al., 2006). Backscatter dataset harmonization may be used to avoid the fragmentation of environmental variables for the continuous prediction of benthic species distributions. Furthermore, the harmonized layer can be utilized for ground-truth sample design prior to species or habitat distribution modelling. An efficient approach to acquiring ground truth across a range of benthic habitats is to target distinct environments *a priori*. This is often accomplished using stratified sampling, where ground truth sites are allocated across the range of environmental data values expected to influence the distribution of interest. Backscatter is a common proxy used to represent substrate characteristics for such designs, yet values of the backscatter layer must represent substrate properties

consistently for the stratification to be meaningful. Backscatter harmonization methods may be useful for sample planning where discontinuous multisource datasets hinder the targeting of distinct substrate types, and data binning for the stratification should make the harmonized layer more robust to minor differences in substrate-specific response between datasets.

5. Conclusions

Here, benthoscape classification results produced using a bulk shift-harmonized backscatter layer were compared to those from independent, non-harmonized multisource layers. Findings suggested that the harmonization provides several benefits for benthoscape classification when using disparate backscatter data sources, including a more efficient analysis, increased availability of training data, and a corresponding increase in statistical discrimination between classes. We also emphasize risks associated with the use of a harmonized multisource backscatter layer though – namely, the potential for error introduction from the detection of different bottom properties by different sonars (e.g., due to differences in operating frequency). Where individual datasets have adequate ground truth, treating them independently requires fewer assumptions regarding compatibility of the data.

The proliferation of seabed mapping activities over the past two decades has produced a wealth of data describing the seabed environment, yet new challenges on the compatibility of compiled multisource datasets require novel solutions. Compilation of acoustic backscatter is especially challenging, given that recorded data values can be particular to a given sonar model, set of operating specifications, or even a specific survey. Despite these difficulties, and given the high cost of marine surveying, it is critical to develop robust methods for utilizing seabed mapping datasets collected under a variety of circumstances. While particularly relevant for legacy datasets, pragmatic approaches for maximizing the use of

contemporary survey data are also necessary. As demonstrated in this study, the harmonization of disparate, uncalibrated backscatter datasets is feasible, and can be useful for multisource thematic seafloor mapping applications.

Funding

B. Misiuk was supported by an International Postdoctoral Fellowship awarded through the Ocean Frontier Institute/Canada First Research Excellence Fund.

Appendix A. Bulk shift model residuals

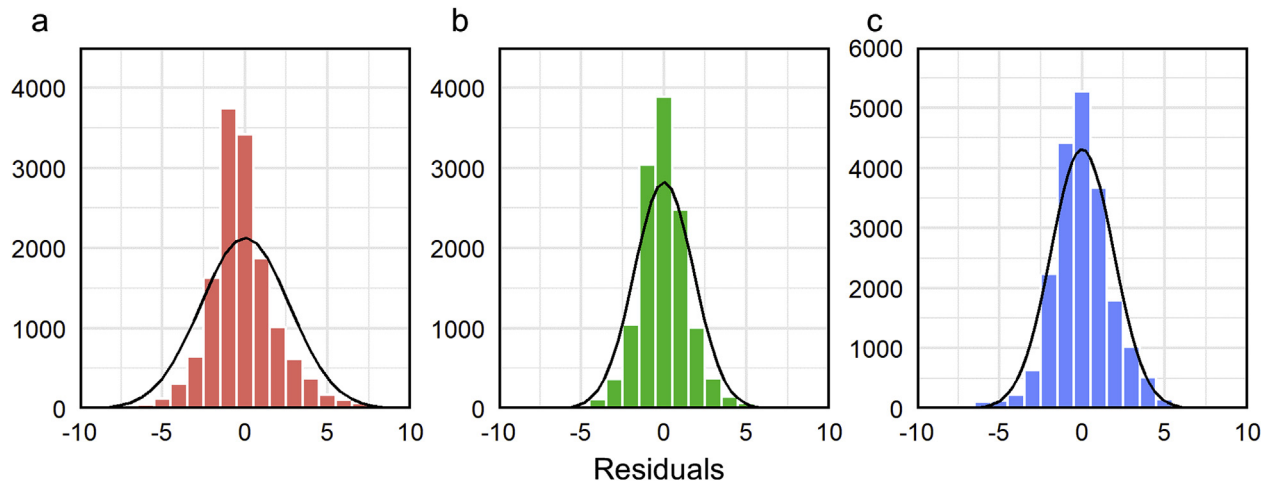


Fig. A1. Bulk shift model residual histograms and theoretical normal curves for (a) EM3002, (b) EM1002, and (c) Reson7111 backscatter datasets.

References

- Boswarva, K., Butters, A., Fox, C.J., Howe, J.A., Narayanaswamy, B., 2018. Improving marine habitat mapping using high-resolution acoustic data; a predictive habitat map for the Firth of Lorn, Scotland. *Contin. Shelf Res.* 168, 39–47. <https://doi.org/10.1016/j.csr.2018.09.005>.
- Brown, C.J., Blondel, P., 2009. Developments in the application of multibeam sonar backscatter for seafloor habitat mapping. *Appl. Acoust.* 70, 1242–1247. <https://doi.org/10.1016/j.apacoust.2008.08.004>.
- Brown, C.J., Smith, S.J., Lawton, P., Anderson, J.T., 2011. Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuar. Coast Shelf Sci.* 92, 502–520. <https://doi.org/10.1016/j.ecss.2011.02.007>.
- Brown, C.J., Sameoto, J.A., Smith, S.J., 2012. Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *J. Sea Res.* 72, 1–13. <https://doi.org/10.1016/j.seares.2012.04.009>.
- Brown, C.J., Beaudoin, J., Brissette, M., Gazzola, V., 2019. Multispectral multibeam echo sounder backscatter as a tool for improved seafloor characterization. *Geosciences* 9, 126. <https://doi.org/10.3390/geosciences9030126>.
- De Falco, G., Tonielli, R., Di Martino, G., Innangi, S., Simeone, S., Parnum, I.M., 2010. Relationships between multibeam backscatter, sediment grain size and *Posidonia oceanica* seagrass distribution. *Contin. Shelf Res.* 30, 1941–1950. <https://doi.org/10.1016/j.csr.2010.09.006>.
- Diesing, M., Green, S.L., Stephens, D., Lark, R.M., Stewart, H.A., Dove, D., 2014. Mapping seabed sediments: comparison of manual, geostatistical, object-based image analysis and machine learning approaches. *Contin. Shelf Res.* 84, 107–119. <https://doi.org/10.1016/j.csr.2014.05.004>.
- Ferrini, V.L., Flood, R.D., 2006. The effects of fine-scale surface roughness and grain size on 300 kHz multibeam backscatter intensity in sandy marine sedimentary environments. *Mar. Geol.* 228, 153–172. <https://doi.org/10.1016/j.margeo.2005.11.010>.
- Fonseca, L., Brown, C.J., Calder, B., Mayer, L., Rzhonov, Y., 2009. Angular range analysis of acoustic themes from Stanton Banks Ireland: a link between visual interpretation and multibeam echosounder angular signatures. *Appl. Acoust.* 70, 1298–1304. <https://doi.org/10.1016/j.apacoust.2008.09.008>.
- Gaida, T.C., Tengku Ali, T.A., Snellen, M., Amiri-Simkooei, A., van Dijk, T.A.G.P., Simons, D.G., 2018. A multispectral Bayesian classification method for increased acoustic discrimination of seabed sediments using multi-frequency multibeam backscatter data. *Geosciences* 8, 455. <https://doi.org/10.3390/geosciences8120455>.
- Gaida, T.C., Mohammadloo, T.H., Snellen, M., Simons, D.G., 2019. Mapping the seabed and shallow subsurface with multi-frequency multibeam echosounders. *Rem. Sens.* 12, 52. <https://doi.org/10.3390/rs12010052>.
- Hamilton, E.L., 1972. Compressional-wave attenuation in marine sediments. *Geophysics* 37, 620–646. <https://doi.org/10.1190/1.1440287>.
- Heaton, J.L., Rice, G., Weber, T.C., 2017. An extended surface target for high-frequency multibeam echo sounder calibration. *J. Acoust. Soc. Am.* 141, EL388–EL394. <https://doi.org/10.1121/1.4980006>.
- Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29, 773–785. <https://doi.org/10.1111/j.0906-7590.2006.04700.x>.
- Hughes Clarke, J.E., Mayer, L.A., Wells, D.E., 1996. Shallow-water imaging multibeam sonars: a new tool for investigating seafloor processes in the coastal zone and on the continental shelf. *Mar. Geophys. Res.* 18, 607–629. <https://doi.org/10.1007/BF00313877>.
- Innangi, S., Tonielli, R., Romagnoli, C., Budillon, F., Di Martino, G., Innangi, M., Laterza, R., Le Bas, T., Lo Iacono, C., 2019. Seabed mapping in the Pelagie Islands marine protected area (Sicily Channel, southern Mediterranean) using remote sensing object based image analysis (RSOBIA). *Mar. Geophys. Res.* 40, 333–355. <https://doi.org/10.1007/s11001-018-9371-6>.
- Janowski, L., Madricardo, F., Fogarin, S., Kruss, A., Molinaroli, E., Kubowicz-Grajewska, A., Tegowski, J., 2020. Spatial and temporal changes of tidal inlet using object-based image analysis of multibeam echosounder measurements: a Case from the Lagoon of Venice, Italy. *Rem. Sens.* 12, 2117. <https://doi.org/10.3390/rs12132117>.
- Kostylev, V.E., Courtney, R.C., Robert, G., Todd, B.J., 2003. Stock evaluation of giant scallop (*Placopecten magellanicus*) using high-resolution acoustics for seabed mapping. *Fish. Res.* 60, 479–492. [https://doi.org/10.1016/S0165-7836\(02\)00100-5](https://doi.org/10.1016/S0165-7836(02)00100-5).
- Lacharité, M., Brown, C.J., Gazzola, V., 2018. Multisource multibeam backscatter data: developing a strategy for the production of benthic habitat maps using semi-automated seafloor classification methods. *Mar. Geophys. Res.* 39, 307–322. <https://doi.org/10.1007/s11001-017-9331-6>.

- Lamarche, G., Lurton, X., 2018. Recommendations for improved and coherent acquisition and processing of backscatter data from seafloor-mapping sonars. *Mar. Geophys. Res.* 39, 5–22. <https://doi.org/10.1007/s11001-017-9315-6>.
- Lamarche, G., Lurton, X., Verdier, A.-L., Augustin, J.-M., 2011. Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—application to Cook Strait, New Zealand. *Continental Shelf Res.* 31, S93–S109. <https://doi.org/10.1016/j.csr.2010.06.001>.
- Lucieer, V.L., 2008. Object-oriented classification of sidescan sonar data for mapping benthic marine habitats. *Int. J. Rem. Sens.* 29, 905–921. <https://doi.org/10.1080/01431160701311309>.
- Lucieer, V., Roche, M., Degrendele, K., Malik, M., Dolan, M., Lamarche, G., 2018. User expectations for multibeam echo sounders backscatter strength data-looking back into the future. *Mar. Geophys. Res.* 39, 23–40. <https://doi.org/10.1007/s11001-017-9316-5>.
- Lurton, X., 2010. In: *An Introduction to Underwater Acoustics: Principles and Applications*, third ed. Springer-Verlag, Berlin Heidelberg, Germany.
- Lurton, X., Lamarche, G. (Eds.), 2015. *Backscatter Measurements by Seafloor-Mapping Sonars: Guidelines and Recommendations*. GeoHab Backscatter Working Group.
- Malik, M., 2019. Sources and impacts of bottom slope uncertainty on estimation of seafloor backscatter from swath sonars. *Geosciences* 9, 183. <https://doi.org/10.3390/geosciences9040183>.
- Malik, M., Lurton, X., Mayer, L., 2018. A framework to quantify uncertainties of seafloor backscatter from swath mapping echosounders. *Mar. Geophys. Res.* 39, 151–168. <https://doi.org/10.1007/s11001-018-9346-7>.
- Misiuk, B., Lecours, V., Bell, T., 2018. A multiscale approach to mapping seabed sediments. *PLoS One* 13, e0193647. <https://doi.org/10.1371/journal.pone.0193647>.
- Misiuk, B., Brown, C.J., Robert, K., Lacharité, M., 2020. Harmonizing multi-source sonar backscatter datasets for seabed mapping using bulk shift approaches. *Rem. Sens.* 12, 601. <https://doi.org/10.3390/rs12040601>.
- Monterea-Gavazzi, G., Roche, M., Lurton, X., Degrendele, K., Terseleer, N., Van Lancker, V., 2018. Seafloor change detection using multibeam echosounder backscatter: case study on the Belgian part of the North Sea. *Mar. Geophys. Res.* 39, 229–247. <https://doi.org/10.1007/s11001-017-9323-6>.
- Monterea-Gavazzi, G., Roche, M., Degrendele, K., Lurton, X., Terseleer, N., Baeye, M., Francken, F., Van Lancker, V., 2019. Insights into the short-term tidal variability of multibeam backscatter from field experiments on different seafloor types. *Geosciences* 9, 34. <https://doi.org/10.3390/geosciences9010034>.
- Passaro, S., Barra, M., Saggiomo, R., Di Giacomo, S., Leotta, A., Uhlen, H., Mazzola, S., 2013. Multi-resolution morpho-bathymetric survey results at the Pozzuoli–Baia underwater archaeological site (Naples, Italy). *J. Archaeol. Sci.* 40, 1268–1278. <https://doi.org/10.1016/j.jas.2012.09.035>.
- Plets, R., Quinn, R., Forsythe, W., Westley, K., Bell, T., Benetti, S., McGrath, F., Robinson, R., 2011. Using multibeam echo-sounder data to identify shipwreck sites: archaeological assessment of the Joint Irish Bathymetric Survey data. *Int. J. Naut. Charting.* 40, 87–98. <https://doi.org/10.1111/j.1095-9270.2010.00271.x>.
- Plets, R., Clements, A., Quinn, R., Strong, J., Breen, J., Edwards, H., 2012. Marine substratum map of the causeway coast, northern Ireland. *J. Maps* 8, 1–13. <https://doi.org/10.1080/17445647.2012.661957>.
- R Core Team, 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Roche, M., Degrendele, K., Vrignaud, C., Loyer, S., Le Bas, T., Augustin, J.-M., Lurton, X., 2018. Control of the repeatability of high frequency multibeam echosounder backscatter by using natural reference areas. *Mar. Geophys. Res.* 39, 89–104. <https://doi.org/10.1007/s11001-018-9343-x>.
- Snellen, M., Gaida, T.C., Koop, L., Alevizos, E., Simons, D.G., 2019. Performance of multibeam echosounder backscatter-based classification for monitoring sediment distributions using multitemporal large-scale ocean data sets. *IEEE J. Ocean. Eng.* 44, 142–155. <https://doi.org/10.1109/JOE.2018.2791878>.
- Stephens, D., Diesing, M., 2014. A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PLoS One* 9, e93950. <https://doi.org/10.1371/journal.pone.0093950>.
- Thorsnes, T., Chand, S., Brunstad, H., Lepland, A., Lågstad, P., 2019. Strategy for detection and high-resolution characterization of authigenic carbonate cold seep habitats using ships and autonomous underwater vehicles on glacially influenced terrain. *Front. Mar. Sci.* 6, 708. <https://doi.org/10.3389/fmars.2019.00708>.
- Todd, B.J., Fader, G.B.J., Courtney, R.C., Pickrill, R.A., 1999. Quaternary geology and surficial sediment processes, Browns Bank, Scotian Shelf, based on multibeam bathymetry. *Mar. Geol.* 162, 165–214. [https://doi.org/10.1016/S0025-3227\(99\)00092-4](https://doi.org/10.1016/S0025-3227(99)00092-4).
- van Rein, H., Brown, C.J., Quinn, R., Breen, J., Schoeman, D., 2011. An evaluation of acoustic seabed classification techniques for marine biotope monitoring over broad-scales (>1 km²) and meso-scales (10 m² – 1 km²). *Estuar. Coast Shelf Sci.* 93, 336–349. <https://doi.org/10.1016/j.ecss.2011.04.011>.
- Walbridge, S., Slocum, N., Pobuda, M., Wright, D.J., 2018. Unified geomorphological analysis workflows with benthic terrain modeler. *Geosciences* 8, 94. <https://doi.org/10.3390/geosciences8030094>.
- Weber, T.C., Rice, G., Smith, M., 2018. Toward a standard line for use in multibeam echo sounder calibration. *Mar. Geophys. Res.* 39, 75–87. <https://doi.org/10.1007/s11001-017-9334-3>.